UNIVERSITY OF BIRMINGHAM

Centre for Language Studies

MA TEFL/TESL

Module 6 – Corpus Linguistics

March 2009

"A Corpus Study of 'Cup of [tea]' and 'Mug of [tea]'"

By

Brett Laybutt

CL/06/01

Take a small number of words or phrases (between 2 and 5) and do a corpus study to show how they are used in similar or different ways. Choose words/phrases which are interesting in some way e.g. your students often confuse them; they cause problems for translators working with a specific language; you yourself have difficulty deciding when to use one or the other. Examples of words/phrases which have been studied in the past include: *between* and *through*; *immense, enormous* and *massive*; *reason to* and *reason for*; *on the other hand* and *on the contrary*. (You should not repeat these studies, which are mentioned as examples, but should choose different sets of words. You may choose words from a language other than English, if you have an available corpus, but make sure that a monolingual English speaker can understand your argument.)

(Approx. 4314 words, excluding footnotes, references, figures, tables, appendices & long quotations.)

Table of Contents

1.0 Introduction

The following paper arose from an English as a Foreign Language (EFL) classroom discussion on loanwords from English into Japanese, in particular the word マグカップ (*magucuppu*). As this loanword is apparently a mix of the English 'mug' and 'cup' the (all Japanese) students were interested as to the differences between these two items. As will be shown below, these differences have been examined within the field of formal semantics from a number of perspectives. However, the forms 'cup of [tea]' and 'mug of [tea]' have largely been overlooked or treated as merely a grammatical extension of the core words. This paper will present a corpus investigation into 'cup of [tea]' and 'mug of [tea]' and demonstrate that the two items are, in fact, distinguished by patterns of linguistic usage as opposed to purely semantic or syntactic constraints. Sections 3.0 and 4.0 will present the results of this investigation, and Section 5.0 will then discuss the implications of these results for corpus linguistics and EFL classes. First, however, Section 2.0 will provide some background on semantics and corpus linguistics.

2.0 Literature Review

The following section will be divided into two sections. The first section will outline some semantic approaches to explicating the core 'cup' and 'mug', and their relation to 'cup of [tea]' and 'mug of [tea]'. The second will outline corpus linguistics and its contribution to this field and the approaches that will then be used in the current study.

2.1. Semantic definitions of 'cup' and 'mug'

There have been several attempts within the field of semantics to explicate the features that differentiate 'cup' and 'mug', a distinction of "notorious difficulty" (Carter, 1998, p. 19). One of the first, and most influential, was Labov's (2004) original 1975 experiment in which subjects were shown pictures of varying indeterminacy (Appendix 1) and asked to label them. From this, Labov was able to come up with a mathematical definition of 'cup' as:

Figure 1: Labov's (2004) definition of 'cup'

The term *cup* is used to denote round containers with a ratio of depth to width of $1 \pm r$ where $r \leq r_b$, and $r_b = \alpha_1 + \alpha_2 + \ldots \alpha_\upsilon$ and α_1 is a positive quality when the feature i is present and 0 otherwise.

feature   1 = with one handle
          2 = made of opaque vitreous material
          3 = used for consumption of food
          4 = used for the consumption of liquid food
          5 = used for consumption of hot liquid food
          6 = with a saucer
          7 = tapering
          8 = circular in cross-section

Cup is used variably to denote such containers with ratios width to depth $1 \pm r$ where $r_b \leq r \leq r_1$ with a probability of $r_1 - r/r_t - r_b$. The quantity $1 \pm r_b$ expresses the distance from the modal value of width to height.

(Labov, 2004, p. 86)

While this definition explicates 'cup', Labov saw a 'mug' as merely a type of cup used for coffee (Labov, 2004). There is also little indication of how this core definition is affected by grammatical contexts in the form of 'cup of [tea]'.

One of the most detailed attempts at explicating the two core items 'cup' and 'mug' was Wierzbicka (1985) and Goddard's (1998) Natural Semantic Metalanguage (NSM) approach. NSM arose from Leibniz's concept that "one cannot define all words" (Wierzbicka, 1996, p. 9). As such, a core of semantic primitives (Appendix 2) can be used to define other words, such as 'cup' in Figure 2:

Figure 2: NSM definition of 'cup'

A *cup* =

    a.   a kind of thing
        people us them for drinking tea or coffee from
    b.   they use them like this:
            someone pours the tea or coffee into it from above while it rests on something
            after this a person picks it up with the fingers of one hand
            raises it up to the mouth so part of the top touches the lower lip
            tips it towards the mouth for a short time so a little tea or coffee moves down inside the mouth
            then puts it down on the same thing as before
            after this, the same person can do all this a few more times
    c.   people make things of this kind so they can use them like this
        because of this:
            they are made of something hard and smooth and thin
            they are round and open at the top
            the bottom is round and flat, it is smaller than the top
            often they have a thin rounded part sticking out from one side
            which a person can hold between the fingers and thumb of one hand
    d.   often someone drinking from one of these things uses another kind of thing to put it down on
        the person drinking puts it down in the middle of this other kind of thing
        these other things are made of the same hard, smooth stuff
        they are round and the have an outside edge which is higher than the middle
        so that if any liquid comes down onto them, it stays there

                                    (Goddard, 1998, p. 233)

Or 'mug' in Figure 3:

Figure 3: NSM definition of 'mug'

A *mug* =

    a.   a kind of thing
          people use this kind of thing for drinking something hot from
    b.   they use it like this:
               someone pours the hot drink into it from above
               a person raises it to the mouth so part of it touches the lip
               then tips the top part towards the mouth so a little liquid moves down inside the mouth
          a person can easily hold a thing of this kind with one hand for a long time
          because of this, a person can drink from it for a long time without having to put it down
    c.   people make things of this kind so they can use them like this
          because of this:
               they are made of something hard and smooth
               they are round and open at the top
               the bottom is round and flat
               they have a rounded part sticking out from one side which a person can put the fingers through
               they are big enough to have as much in them as a person would want to drink at one time

(Goddard, 1998, p. 236)

Within this conception, however, 'cup of coffee', for example, is seen as a separate lexicalised expression (Goddard, 1998). It is thus not clear what the relationship between 'cup of [tea]' and 'mug of [tea]' might be, and whether 'mug of [tea]' can also be considered a lexicalised phrase or the result of a grammatical process. This definition also does not give any indication of the differing senses of the two core words, information that would be important for any learners of a foreign language.

Turning to lexicographical definitions, the differing senses of 'cup' and 'mug' become clearer. The Oxford Advanced Learner's Dictionary (2005) gives a definition of 'cup' (Appendix 3a) with a total of ten senses for the noun, and 'mug' (Appendix 3b) a total of four. From these dictionary definitions, it could be taken that 'cup of [tea]' is confined to two senses, that of

"contents" and "measure" and 'mug of [tea]' is confined to just one sense, that of "contents". As will be seen below, however, the relation between the items and the senses is more complicated than this would suggest.

From these semantic definitions then it would seem that while 'cup' and 'mug' can be clearly explicated into their core components and senses and do, in fact, have separate prototypical meanings as suggested by Taylor (2002), the differences between 'cup of [tea]' and 'mug of [tea]' are not so clear and are often seen as somewhat synonymous grammatical constructions. It also appears, however, that purely semantic distinctions are not adequate to separate them. The next section will look at the role corpus linguistics might play in explicating the two terms.

2.2. Corpus linguistics

The following section will provide some background to corpus linguistics and the role of corpus linguistics in explicating near-synonymous words. This will be followed by an overview of how this role can be applied within the EFL classroom.

2.2.1. Background to corpus linguistics

A corpus can be generally defined as "a collection of texts in an electronic database" (Kennedy, 1998, p. 3). While collections of texts and subsequent concordances had previously been used in various areas (Kennedy, 1998), it was not until the 1950s that computing power was great enough to allow the storage of large amounts of text for linguistic exploration. Within neo-

Chomskian linguistics prevalent at the time, these first corpora were met with "massive indifference if not outright hostility" (Kennedy, 1998, p. 23) as it was felt that investigations of linguistic performance held no particular contribution for understanding of language (Halliday, 1991). These 1st generation corpora were relatively small but laid the groundwork for establishing a framework of principles for subsequent corpus design.

The 2nd generation of corpora, from the 1990s, were specifically designed to be as representative as possible of natural language and large enough to be able to make generalised statements about language. These so-called 'mega-corpora' were also revolutionary in that they were specifically linked to commercial activity (Kennedy, 1998); the Cobuild dictionary, for example, arose out of collaboration between Collins publishers and Birmingham University and was the first to utilise corpus data in its compilation, a situation now mostly taken for granted (Rundell, 2008).

2.2.2. Types of corpora

There are generally recognised to be eight differing types of corpora (Hunston, 2002; Kennedy, 1998), outlined in Appendix 4. These range from very large general and monitor corpora to much smaller specialised ones, all of which have various advantages and disadvantages depending on the study in question. The following study will use the Bank of English monitor corpus, described in more detail in Section 3.2 below. The advantage of using the large corpus is that its very size allows a better calculation of frequency and collocation information (Sinclair, 1991) and also a better indication of subtle patterns associated with particular items (Hunston, 2002) and semantic prosody (Hunston, 2007).

2.2.3. Uses of corpus linguistics

The number and variation of the types of corpora reflects the fact that corpora can be used to investigate language in a number of different ways and for a number of different purposes. It has become accepted that corpus investigations can reveal insights into language use that intuition may overlook (McEnery, Xiao, & Tono, 2006). One of the main uses of corpora is to investigate phenomena such as phraseology and pattern grammar (Hunston & Francis, 2000; Hunston, 2002), where the study of concordances may highlight recurrent patterns in wider contexts and the relation between these patterns and meaning (Hunston, 2002; Sinclair, 1991). Corpora may also be used to investigate the collocation and semantic prosody (Hunston, 2007; Partington, 1998; Sinclair, 1991) surrounding particular words, which may also be affected by register variation (Biber, 1995). It also has uses within language teaching (Gavioli, 2005) and lexicology (Rundell, 2008).

One of the most useful applications of corpus linguistics is in the investigation of near-synonymy. One of the most exemplary was Kennedy's (1991) study of the behaviour of *between* and *through* from the perspective of collocation and semantic function. Some other studies include Partington's (1998) discussion of *sheer*, *pure* and *absolute* in terms of collocation and prosody, or Tognini-Bonelli's (2001) studies of *largely* and *broadly* or *tall* and *large*. For EFL teaching in particular, Johns (1991) has demonstrated the value of using concordances to highlight near-synonyms such as *persuade* and *convince*. More recently, corpus studies have also found application within pragmatics and discourse studies in order to "provide information about social and textual factors that influence language choice" (Conrad, 2002, p. 75). The following study

will attempt to demonstrate the differences between 'cup of [tea]' and 'mug of [tea]' from some of these perspectives and methodologies.

2.2.4. Corpus linguistics and language teaching

While there is little doubt that the findings of corpus studies have shed new light on language, there is still some debate as to the relative value of corpora within foreign language teaching (McEnery, Xiao, & Tono, 2006). Corpus data has increasingly revealed the interrelated nature of lexis and grammar leading, for example, to Halliday's (2004) concept of a 'lexico-grammar', and its probabilistic nature (Halliday, 1991). How this data may be applicable within the classroom, however, has not found as widespread acceptance. Suggestions have ranged from the teaching of prefabricated chunks (Nattinger & DeCarrico, 1992), Data Driven Learning (Johns, 1991) using actual corpus data, lexical phrases (Willis, 1990) or through collocation (Lewis, 2001). All of these have their various merits and drawbacks yet it still remains the case that most lexical approaches to language teaching lack a coherent methodology (Richards & Rodgers, 2001). Section 5 will discuss some of these issues further in light of the following study.

3.0 Methodology

The following section will outline the methodology to be used in this study. The aims of the study and source of data will first be described followed by the procedure to investigate differences between 'cup of [tea]' and 'mug of [tea]'.

3.1. Aims of the current study

In terms of formal semantics the differences between 'cup' and 'mug' have been explicated from a number of differing points of view, as was described in Section 2.0. However, whether semantics can account for differences between 'cup of [tea]' and 'mug of [tea]' is not as clear and corpus linguistics may thus be better equipped to separate them. The following study, then, will present a corpus study of the phrases 'cup of [tea]' and 'mug of [tea]' and investigate the linguistic usage of the two phrases in terms of collocation, phraseology and pragmatic and discourse patterns.

3.2. The 'Bank of English' Corpus

The data comes from the Bank of English (BoE) corpus jointly owned by HarperCollins Publishers and the University of Birmingham. The BoE is a monitor corpus designed to be as large as possible and to continually gather new data (Sinclair, 1991; Kennedy, 1998), reflecting Sinclair's (1991) belief that large amounts of data are necessary for significant statistical analysis to be conducted. In 2007 the corpus stands at 450 million words.

3.3. Procedure

For the following study the main procedural problem that must be overcome before any detailed investigation can begin is large amounts of data. A search for 'cup of [tea]' on the BoE reveals a total of 5298 lines, while 'mug of [tea]' reveals 355. It is generally considered that for looking at general patterns roughly 100 lines is at best manageable (Hunston, 2002). The procedure will thus follow Hunston's (2002, p. 52) suggestion of "hypothesis testing" for coping with large amounts of data. First, from the search of 'cup of [tea]', a random 100 lines will be selected and investigated for significant patterns in the areas of collocation, prosody and larger patterns of phraseology. This will be continued until further investigation does not reveal any interesting patterns. This will then be repeated for 'mug of [tea]'. Once this process has been completed 'cup of [tea]' and 'mug of [tea]' can compared for differences in usage. The corpus data will then be examined for differences between the two items in terms of pragmatic usage and discourse functions. A note that must be made at this stage is that, due to the word limitations of this paper, the following can only be seen as an exploratory, rather than in depth, study.

4.0 Results

The following section will present the results of a corpus study of 'cup of [tea]' and 'mug of [tea]'. First, the two items will be examined in terms of collocation and semantic prosody. Next, some differences in phraseology will be demonstrated, followed by pragmatic and discourse functions. All examples are taken from the Bank of English corpus.

4.1. Collocation and semantic prosody of 'cup of [tea]' and 'mug of [tea]'

Turning first to items which collocate with 'cup of [tea]' and 'mug of [tea]' in the sense of "contents" we can see that there is similarity between them, in that the ratio of the top two collocates, *tea* and *coffee*, is roughly the same at 2:1. However, these two collocates do seem significantly higher by t-score compared to other collocates for 'cup of [tea]' than those for 'mug of [tea]'.

Table 1: Collocation with 'cup of [tea]' and 'mug of [tea]' by t-score (span = 0:1)

Contents	'cup of ?'	n	t-score	'mug of ?'	n	t-score
	tea	2184	46.731189	tea	102	10.098489
	coffee	1234	35.125476	coffee	59	7.6797757
	water	71	8.3698892	cocoa	13	3.6053917
	cocoa	34	5.8300099	beer	10	3.1606097
	soup	22	4.6856551	water	6	2.4292097
	milk	22	4.6658525	ovaltine	6	2.4494690
	chocolate	12	3.4599030	horlicks	5	2.2360594
	nescafe	10	3.1619720			
	wine	10	3.1369709			
	herbal	10	3.1615524			
	espresso	9	2.9998390			
	cappuccino	8	2.8284056			

While they both may refer to liquid contents it is also the case that 'mug of [tea]' is not necessarily liquid but may be merely the sense of a container, for example:

the mug of spoons

'Cup of [tea]', however, may also be used metaphorically, for example:

Cup of cheer
a cup of holiday cheer
the bitter cup of existential woe
the unsatisfying egg cup of information

While at first glance they both appear to refer to the sense of "contents" as in the dictionary definitions, looking next at verbs associated with the two items Table 2 shows that those verbs concerned with dealing with liquid, such as *drank*, *poured* or *sipped*, are more likely collocates of 'cup of [tea]', whereas collocates of 'mug of [tea]' seem to refer more to the object itself; *drank*, for instance, is not among the list of collocates:

Table 2: Collocation of verbs with 'cup of [tea]' and 'mug of [tea]' by t-score (span = 4:4)

Sense	'cup of'	n	t-score	'mug of'	n	t-score
"liquid"	had	121	10.373396	made	6	2.402322
	made	107	10.237491	poured	5	2.234147
	poured	71	8.421286	had	5	1.913061
	offered	49	6.985394	filled	4	1.957693
	drank	28	5.284294			
	finished	12	3.435024			
	sipped	10	3.158880			
	asked	10	2.998968			
	accepted	9	2.966846			
"container"	put	30	5.339668	handed	12	3.462661
	handed	17	4.111557	set	5	2.216847
	gave	16	3.920848	picked	6	2.445286
	held	11	3.246039	held	4	1.987734
				brought	4	1.985634
				put	4	1.960524
				gave	3	1.712896
				grabbed	3	1.720273

Both items seem to make use of 'V-ing' within their phrasing (which will be discussed further in Section 4.3), yet here as well there are differences between them. The top collocates, shown in Table 3 below, indicate that, again, 'mug of [tea]' is more likely to be referred to in the sense of an object, with collocates such as *clutching* and *nursing* which are not as significant for 'cup of [tea]', whereas 'cup of [tea]' is more concerned with the sense of liquid contents, with items such as *having*, *making* or *pouring* having a much higher t-score than those for 'mug of [tea]':

Table 3: Comparison of V-ing collocation by t-score (span = 4:4)

V-ing	'cup of'	n	t-score	'mug of'	n	t-score
	having	71	8.3792505	sipping	9	2.9998083
	making	56	7.4428696	clutching	9	2.9995720
	drinking	53	7.2728562	drinking	5	2.2335932
	sipping	21	4.5813775	making	4	1.9841423
	holding	14	3.7129626	having	4	1.9792948
	pouring	20	4.4684247	holding	3	1.7255553
	sitting	19	4.3211560	nursing	3	1.7309955

Comparing semantic prosody by collocation of the two terms, it appears that the prosody of 'mug of [tea]' is simply in terms of physical properties, with collocates such as *warm*, *large* and *cold* having similar t-scores, whereas 'cup of [tea]' describes an evaluative response to the drink through top collocates such as *nice*, *decent*, *enjoy* and *perfect*, as shown in Table 4:

Table 4: Comparison of prosody by t-score (span = 4:4)

Prosody	'cup of'	n	t-score	'mug'	n	t-score
	hot	126	11.21132	hot	27	5.1930637
	nice	85	9.2028646	warm	7	2.6417570
	strong	51	7.1156191	nice	6	2.4429113
	warm	30	5.4588136	instant	7	2.6443674
	enjoy	29	5.3711109	black	9	2.9871199
	decent	25	4.9940733	large	7	2.6297197
	perfect	20	4.4560093	cold	5	2.2266952

In summary, it appears from the collocation and prosody then that 'mug of [tea]' more likely refers to the first sense, "container", while 'cup of [tea]' refers more to the second, as "contents". The former is more likely to be physically described while the latter may be expressed in evaluative terms. The next section will look at larger patterns of phraseology.

4.2. Phraseology of 'cup of [tea]' and 'mug of [tea]'

As mentioned in Section 4.2, one initial finding that could be made is that within the BoE 'mug of [tea]' often appears to occur in the sense of "container" and be physically described in the pattern 'a ADJ-physical mug of [tea]':

sipping from a <u>large</u> mug of tea
he grasped a <u>large</u> mug of tea
nursing an <u>outsize</u> mug of tea
along with a <u>plastic</u> mug of tea
clutching a <u>plastic</u> mug of lager
his hands round a <u>tin</u> mug of coffee
A <u>tin</u> mug of coffee

'Cup of [tea]', on the other hand, is more likely to describe an evaluative response in the pattern 'a ADJ-evaluative cup of [tea]':

I can make a <u>decent</u> cup of tea
He'd like a <u>decent</u> cup of coffee
enjoy a <u>delicious</u> cup of Douwe Egberts
a <u>drinkable</u> cup of tea
a <u>unique and elegant</u> cup of tea
an <u>excellent</u> cup of coffee
an <u>expensive</u> cup of coffee
a <u>damn fine</u> cup of tea
a <u>good</u> cup of coffee
a <u>good</u> cup of tea
a <u>nice</u> cup of chocolate
a <u>nice</u> cup of tea

14

It also appears, however, that both items most commonly take the form 'a cup of [tea]' with no intervening word between *a* and *cup*. In fact, the ratios of '[+a] cup of' to '[-a] cup of' and '[+a] mug of' to '[-a] mug of' appears to be about the same at approx. 20:11. Yet here too, there are differences in phraseology. In the case of 'a cup of [tea]' we are more likely to find extended clause patterns, such as 'V and V a cup of [tea]':

<div align="center">

come by and have a cup of tea
watch television and have a cup of coffee
get up and fix myself a cup of tea
hang out and enjoy a cup of coffee
go off and make a cup of tea
to go and make a cup of tea
I'll go and make a cup of tea
all go and have a cup of tea together
Come and have a cup of tea
Come and have a cup of coffee
ate a ham sandwich and drank a cup of hot coffee
sit back and enjoy a cup of freshly-brewed Sainsbury's

</div>

This is also in the pattern 'V-ing and V-ing a cup of [tea]':

<div align="center">

going to my room and having a cup of tea
strolling over and purchasing a cup of coffee
reading the paper and drinking a cup of tea
eating sandwiches and drinking a cup of coffee
eating a croissant and drinking a cup of coffee
yawning and pouring himself a cup of coffee

</div>

There is also the pattern 'V-ing V-ing a cup of [tea]', especially with *sitting* or *standing*:

<div align="center">

standing with me having a cup of tea
standing on her porch, drinking a cup of coffee
We were sitting having a cup of tea
I was sitting down, drinking a cup of tea
Kath sitting up in bed sipping a cup of tea
Jupe was sitting up, holding a cup of soup

</div>

15

On the other hand, 'mug of [tea]' tends to be in simple clause types rather than extended patterns:

> I was having a mug of tea
> he grasped a large mug of tea
> She's drinking a mug of instant coffee

It may also occur within extended clauses yet, whereas 'a cup of [tea]' mainly appears to remain in the position of direct object, 'mug of [tea]' becomes instead the object of a prepositional phrase, often following *with*:

> pop stars were tucked up in bed with a mug of cocoa
> sitting at the kitchen table, with a mug of tea
> I followed this with a mug of coffee

From the evidence of the BoE then, it does appear that there are some differences in phraseology between 'cup of [tea]' and 'mug of [tea]'. Phraseology, however, is not the only difference between the two. The next sections will look at differences in pragmatic and discourse usage of 'cup of [tea]' and 'mug of [tea]'.

4.3. Pragmatic functions of 'cup of [tea]' and 'mug of [tea]'

The use of offers and requests also shows lexical differences in between the two, which indicates differing levels of politeness and distance, with 'cup of [tea]' being at a higher level. The word OFFER itself, for example, is commonly used with 'cup of [tea]':

> Caili <u>offers</u> me a cup of barley beer
> They <u>offered</u> us a cup of coffee
> BA would <u>offer</u> us a cup of tea
> almost as if he were <u>offering</u> me a cup of tea
> I <u>offered</u> him a cup of tea
> <u>offered</u> to make a colleague a cup of tea

16

As is ACCEPT:

He <u>accepted</u> a cup of tea and a Marie biscuit
eventually <u>accepting</u> a cup of tea
<u>accepting</u> a cup of Patsy's strong tea
to <u>accept</u> his invitation to a cup of coffee
He debated whether to <u>accept</u> a cup of tea

On the other hand, 'mug of [tea]' does not seemingly require an invitation, reflected by the extensive use of lexical items such as, GIVE, HAND and PASS:

Walsh <u>reached for</u> the mug of vile instant coffee
African woman <u>brought</u> him a mug of sweet tea
They'd <u>given</u> him a mug of tea
he'd be <u>given</u> a mug of tea and a slice of bread
She <u>passed</u> me a mug of coffee
She <u>passed</u> a mug of coffee to Irena
He <u>handed</u> me a mug of steaming, sweet sludge
Thompson <u>handed</u> him a mug of tea

In pragmatics, ellipsis refers to the "absence of a word or words from a structural slot" (Yule, 1996, p. 129). The use of ellipsis implies that the listener can infer what the intended referent is through some shared cultural assumptions (Yule, 1996). 'Cup of [tea]' seems to permit ellipsis to a much greater extent than 'mug of [tea]'. The contraction *cuppa*, for example, is possible whereas *mugga* is not; within the BoE there are 408 examples of the former yet none for the latter. There may be several interpretations for this. It may reflect the fact that, as suggested in Section 4.2 above, 'cup of [tea]' generally refers to liquid and the object may thus be taken for granted, whereas 'mug of [tea]' refers to the object itself and so requires a specific liquid to be mentioned. It may also be a result of differing levels of lexicalisation. Another, pragmatic, interpretation may be that the use of ellipsis is a pragmatic attempt at overcoming the usual formality of 'cup of [tea]' and creating social closeness, which is already present in the use of 'mug of [tea]'. The next section will now go on to discuss discourse function.

4.4. Discourse functions of 'cup of [tea]' and 'mug of [tea]'

As far as discourse is concerned, 'cup of [tea]' often appears to be used as a time marker. For example, looking at collocates with 'cup of [tea]' we can identify several related groupings, such as relating to time:

> her <u>five o'clock</u> cup of tea
> a <u>six o'clock</u> cup of tea
> our <u>evening </u>cup of tea
> his <u>evening</u> cup of sake
> his <u>morning</u> cup of Earl Grey tea
> enjoying a <u>morning</u> cup of tea
> that <u>early morning</u> cup of freshly-ground coffee
> your <u>early morning</u> cup of tea
> a <u>mid-morning</u> cup of soup
> she sips a <u>mid-morning</u> cup of coffee
> the <u>half-time</u> cup of Bovril
> Up for a <u>half-time</u> cup of restorative
> there's <u>no time for</u> a cup of tea

Number:

> it's time for my <u>first</u> cup of coffee
> Helena took her <u>first </u>cup of coffee
> having their <u>first</u> cup of coffee
> I was on my <u>second</u> cup of tea
> pouring a <u>second </u>cup of coffee
> He poured his <u>second </u>cup of coffee
> have a <u>second</u> cup of coffee
> poured himself a <u>third</u> cup of tea
> had just finished his <u>third</u> cup of coffee
> I worked on my <u>third</u> cup of coffee
> We think the <u>fourth</u> cup of wine
> over his <u>fourth</u> cup of coffee
> Stone was on his <u>fourth</u> cup of coffee
> By his <u>last</u> cup of coffee

Completion:

> a <u>half-finished</u> cup of coffee
> A <u>half-finished</u> cup of tea
> she drank <u>the last</u> cup of tea

And habit:

>   the <u>occasional</u> cup of nettle tea
>   an <u>occasional</u> cup of coffee
>   she enjoys the <u>occasional</u> cup of tea
>   for the <u>odd</u> cup of tea
>   I'll have the <u>odd</u> cup of tea
>   her <u>usual</u> cup of Kenco
>   a <u>weekly</u> cup of tea
>   the <u>usual</u> cup of tea

On the other hand, 'mug of [tea]' seems to be used more as a position marker, indicating spatial

relations of location between objects:

>   a steaming mug of coffee <u>on the small table</u>
>   She set a mug of coffee <u>down in front of him</u>
>   The mug of coffee was still <u>on the bench</u>
>   balancing a mug of coffee <u>on the steering wheel</u>
>   a mug of coffee <u>on the small table</u>
>   steaming mug of tea <u>served up on a grubby</u>
>   A steaming mug of tea <u>had appeared at my elbow</u>

Or location of participants:

>   around a mug of tea <u>at her kitchen table</u>
>   <u>in the parlor house kitchen</u> with a mug of steaming coffee
>   Over a mug of coffee <u>in the boardroom</u>
>   clutching a plastic mug of lager <u>behind the stands</u>
>   having a mug of tea <u>by the fire</u>
>   <u>sitting at home</u> with a mug of cocoa
>   having a mug of tea <u>in the office</u>
>   I like to have a mug of coffee <u>in bed</u>
>   <u>tucked up in bed</u> with a mug of cocoa

'Mug of [tea]' also tends to highlight physical relations between participants in discourse

through the use of material process verbs related to movement:

She <u>shoved</u> a mug of coffee <u>in his face</u>
he <u>threw</u> a mug of tea <u>across the office at Matt</u>
Walsh <u>reached for</u> the mug of vile instant coffee
She <u>came into the dining-room</u>, a mug of coffee in each hand
She <u>handed</u> me my mug of tea
I <u>got up and took</u> my mug of tea <u>to the sliders</u>
Sergeant-Major, Phillips, <u>held out</u> a mug of tea <u>to him</u>
She <u>passed</u> a mug of coffee <u>to Irena</u>
I <u>take</u> a mug of tea <u>to the terrace</u>

In terms of participants of discourse, 'cup of [tea]' in the BoE can be used to signal the beginning of conversation, often with *over*:

You see", I <u>said</u>, over a cup of tea
<u>chatting</u> informally over a cup of tea
one day, over a cup of tea, I <u>blurted out</u> the whole
over a cup of tea I <u>said</u>: `My family seem to
together to <u>talk</u> about things over a cup of coffee one afternoon
friendly <u>discussion</u> over a cup of coffee

'Mug of [tea]', on the other hand, seems to signal a discourse boundary. Looking at right-hand collocates in Table 5:

Table 5: Right collocates of 'cup of [tea]' and 'mug of [tea]' by t-score (span = 0:3)

'cup of ?'	n	t-score	'mug of ?'	n	t-score
and	629	23.86529732	and	49	6.52829742
in	188	12.19425488	in	20	3.98474764
or	149	11.81865406	on	9	2.73294544
but	91	8.916086197	to	5	0.84390127
at	73	7.820477009	for	5	1.80751502
when	33	5.248843193			
on	42	5.301001549			
while	29	5.258581638			

we can see that whereas 'mug of [tea]' is followed by a preposition or *and* which might signal the end of a clause, 'cup of [tea]' is also commonly followed by a conjunction or wh-word

suggesting it serves to link discourse between clauses. It appears from the evidence of the Bank of English then that there are some indications of differences between the pragmatic and discourse functions of the two items 'cup of [tea]' and 'mug of [tea]'. Coupled with the analysis of the collocation and phraseology patterns presented above it may be possible to conclude that it is the linguistic behaviour of the two that distinguishes them, rather than any inherent semantic distinction. The next sections will consider the implications of these findings.

5.0 Discussion

The following section will present a discussion of some of the implications of this study for corpus linguistics, especially how it relates to EFL classes. First a summary of the results will be presented and a discussion of their practical implications for EFL classes. This will be followed by a discussion of some of the limitations of the study and some suggestions for further research.

5.1. Summary of results

This study analysed 'cup of [tea]' and 'mug of [tea]' from the perspective of collocation and prosody, phraseology, pragmatic usage and discourse function. From the results above, it appears that 'cup of [tea]' is more likely to be used in the sense of 'contents' whereas 'mug of [tea]' is used for 'container'. 'Cup of [tea]' is also more likely to collocate with evaluative terms than 'mug of [tea]', which more commonly collocates with terms for physical description. There is also some evidence of differences in phraseology, with 'cup of [tea]' seemingly expressed within extended patterns or multiple clauses but 'mug of [tea]' more commonly in simple phrase and clause patterns. Within larger contexts of pragmatic usage and discourse function, 'cup of [tea]' appears to reflect linguistically the more formal social contexts in which it may be found and to highlight in discourse functions associated with time, such as habit and completion. 'Mug of [tea]', on the other hand, seems to reflect more the physical and spatial, as opposed to social, relations between participants in discourse. The next sections will discuss the implications of these findings for corpus linguistics and EFL classes, and suggest some further research. First, however, some of the limitations of this study will be considered.

5.2. Limitations of study

Due to word constraints as mentioned above, this study can only be seen as exploratory at most. While some findings emerged out of the study, these may be due merely to the limitations of corpus evidence as it must be remembered that corpus data only presents evidence of that corpus and absence of data in that corpus does not imply absence in language (Hunston, 2002). The study was also a qualitative one, based on observed patterns. A more detailed quantitative study would also be preferable to confirm if these observations are, in fact, significant. It should also be pointed out that most of the examples in Section 4 above seem to be from written sources yet there may be differences between written and spoken uses of the two items. Taking into account Biber's (2003; 1995) findings of variation between registers one must thus be wary of making generalizations of language use, especially in the areas of discourse and pragmatics.

5.3. Practical implications for EFL classes

While this study has looked at 'cup of [tea]' from the perspective of the core meaning of 'cup' as a drinking vessel, it is also true that the overwhelmingly vast majority of the uses of 'cup' are from the domain of sports. This highlights a problem for language learners in that the most frequent item is not necessarily the most central (Hunston, 2002). The study presented above also suggests that even seemingly simple words can have characteristic patterns of collocation and phraseology associated with them, which might present a problem with, for example, DDL approaches utilising corpus data (Johns, 1991) in that it would be impossible to study every word on a syllabus in this depth. It does suggest, however, the benefits, and possibly necessity, of

presenting these simple vocabulary items through context and not solely through semantic meanings, which also has implications for materials design.

The differences in pragmatic and discourse usage between the two items also indicates the need for learners to be made aware of the socio-cultural assumptions that influence both lexical and grammatical choices at a level beyond that of semantics. While the analysis of pragmatics and discourse generally occurs at the macro-level of text (Conrad, 2002), this study also indicates the value of using corpus studies to reveal pragmatic and discourse function at a micro-level of lexico-grammar. Information such as this, which corpus studies may reveal, would be of real benefit for learners of a language, and there is some evidence that students can benefit from direct instruction of pragmatic functions (Lightbrown & Spada, 2006).

5.4. Suggestions for further research

The study presented here also raises questions as to the nature of the relationship between the core concepts of 'cup' and 'mug' and the phrases 'cup of [tea]' and 'mug of [tea]'. From the semantic definitions outlined in Section 2.1, it would appear that the key difference between 'cup' and 'mug' is one of physical properties of a drinking vessel and the forms 'cup of [tea]' and 'mug of [tea]' are thus taken to be merely a grammatical constructions to express their contents. This view, however, does not fully account for the differences in senses between 'cup of [tea]' and 'mug of [tea]' highlighted by this study. In fact, it may be that while the two appear to have the same form these forms are the result of different processes, with 'cup of [tea]' arising from a process of lexicalisation for the social act of drinking tea, and 'mug of [tea]' resulting from

grammatical or syntactic forces. More research is needed, however, into how these processes occur and in what ways they are acquired. As this study has hopefully demonstrated, the pragmatic and discourse functions are seemingly as important as core semantics in explaining the behaviour of the two items, yet more research is needed, especially within corpus linguistics, into this relationship. This also suggests the need for more research into the nature of L1/L2 lexico-grammar acquisition and whether the same processes occur. More research is also required into how the findings from corpus studies such as this one may be applicable within EFL classrooms.

6.0 Conclusion

This essay has demonstrated that 'cup of [tea]' and 'mug of [tea]' may be differentiated through differences in collocation, phraseology and pragmatic and discourse influences on lexico-grammatical choice, rather than solely through semantics. This suggests that the relationship between the core words and lexicalised items or grammatical constructions is more complex than might at first appear. Corpus studies of the pragmatic and discourse functions of lexico-grammatical items may shed light on this relationship and this area of corpus linguistics may offer a line of research that would be of benefit for EFL learners in particular.

Bibliography

Biber, D. (2003). "Variation among University Spoken and Written Registers: A New Multi-Dimensional Analysis". In P. Leistyna, & C. Meyer (Eds.), *Corpus Analysis: Language Structure and Language Use.* Amsterdam/ New York: Rodopi.

Biber, D. (1995). *Dimensions of Register Variation: a cross-linguistic comparison.* Cambridge: Cambridge University Press.

Carter, R. (1998). *Vocabulary – Applied Linguistic Perspectives* (2nd ed.). New York: Routledge.

Conrad, S. (2002). "Corpus linguistic approaches for discourse analysis". *Annual Review of Applied Linguistics* , 75-94.

Gavioli, L. (2005). *Exploring Corpora for ESP Learning.* Amsterdam: John Benjamins Publishing.

Goddard, C. (1998). *Semantic Analysis.* Oxford: Oxford University Press.

Halliday, M. (1991). "Corpus studies and probabilistic grammar". In K. Aijmer, & B. Altenberg (Eds.), *English Corpus Linguistics.* London and New York: Longman.

Halliday, M. (2004). *An Introduction to Functional Grammar* (3rd (Revised by Matthiessen, C.M.I.M) ed.). London: Hodder Arnold.

Hunston, S. (2002). *Corpora in Applied Linguistics.* Cambridge: Cambridge University Press.

Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics* , 149-268.

Hunston, S., & Francis, G. (2000). *Pattern Grammar - A corpus-driven approach to the lexical grammar of English* . Amsterdam: John Benjamins Publishing.

Johns, T. (1991). "Should you be Persuaded: Two Samples of Data-Driven Learning Materials". *English language research journal* , *4*, 1-16.

Kennedy, G. (1991). "Between and through: The company they keep and the functions they serve". In K. Aijmer, & B. Altenberg (Eds.), *English Corpus Linguistics.* London and New York: Longman.

Kennedy, G. (1998). *An Introduction to Corpus Linguistics.* London: Longman.

Labov, W. (2004). "The Boundaries of Words and their Meanings". In B. Aarts, D. Denison, E. Keizer, & G. Popova (Eds.), *Fuzzy Grammar: A Reader.* Oxford: Oxford University Press.

Lewis, M. (Ed.). (2001). *Teaching Collocation: Further developments in the Lexical Approach.* Hove, England: Language Teaching Publications.

Lightbrown, P., & Spada, N. (2006). *How Languages are Learned* (3rd ed.). Oxford: Oxford University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based Language Studies: An Advanced Resource Book.* New York: Routledge.

Nattinger, J., & DeCarrico, J. (1992). *Lexical Phrases and Language Teaching.* Oxford: Oxford University Press.

Oxford Advanced Learner's Dictionary. (2005). Oxford: Oxford University press.

Partington, A. (1998). *Patterns and Meaning: Using corpora for English language research and teaching.* Amsterdam: John Benjamins.

Richards, J., & Rodgers, T. (2001). *Approaches and Methods in Language Teaching.* Cambridge: Cambridge University Press.

Rundell, M. (2008). "More Than One Way to Skin a Cat: Why Full-sentence Definitions Have Not Been Universally Adopted". In T. Fontenelle (Ed.), *Practical Lexicography: A Reader.* Oxford: Oxford University Press.

Sinclair, J. (1991). *Corpus, Concordance and Collocation.* Oxford: Oxford University Press.

Taylor, J. (2002). "Near synonyms as co-extensive categories: 'high' and 'tall' revisited". *Language Sciences , 25*, 263-284.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work.* ??: John Benjamins Publishing.

Wierzbicka, A. (2006). *English: Meaning and Culture.* Oxford: Oxford University Press.

Wierzbicka, A. (1985). *Lexicography and conceptual analysis.* Ann Arbor: Karoma.

Wierzbicka, A. (1996). *Semantics: Primes and Universals.* Oxford: Oxford University Press.

Willis, D. (1990). *The lexical syllabus: a new approach to language teaching.* London and Glasgow: Collins ELT.

Yule, G. (1996). *Pragmatics.* Oxford: Oxford University Press.

Appendix 1: Labov's (2004) pictures of cups
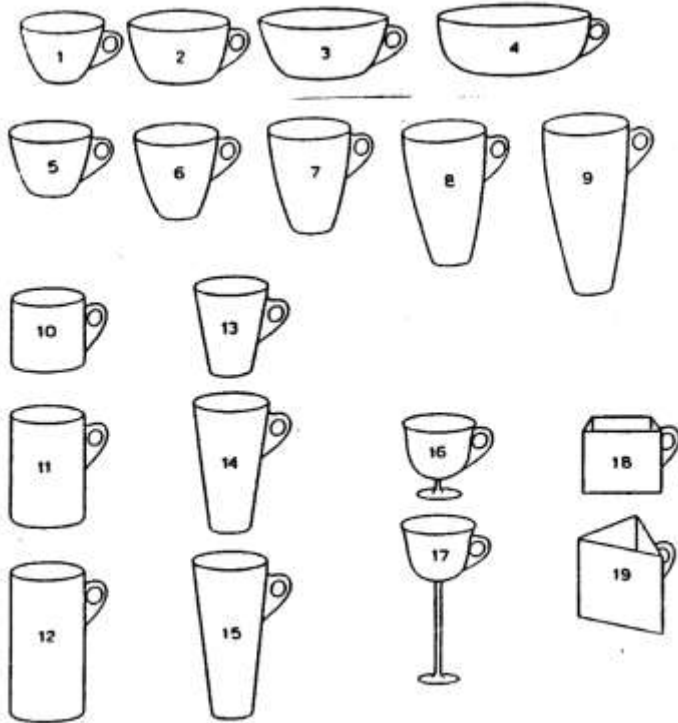
(Labov, 2004)

Appendix 2: Table of semantic primes

Meaning	Semantic Prime
Substantives	I, YOU, SOMEONE/PERSON, SOMETHING/THING, PEOPLE, BODY
Relational substantives	KIND, PART
Determiners	THIS, THE SAME, OTHER,ELSE
Quantifiers	ONE, TWO, SOME, ALL, MUCH/MANY
Evaluators	GOOD, BAD
Descriptors	BIG, SMALL
Mental/experiential predicates	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech	SAY, WORDS, TRUE
Actions, events, movement	DO, HAPPEN, MOVE
Existence and possession	THERE IS/EXIST, HAVE
Life and death	LIVE, DIE
Time	WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT
Space	WHERE/PLACE, BE (SOMEWHERE), HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE, TOUCHING
Logical concepts	NOT, MAYBE, CAN, BECAUSE, IF
Augmentor, intensifier	VERY, MORE
Similarity	LIKE (AS, HOW)

(Wierzbicka, 2006, p. 18)

Appendix 3a: Oxford Advanced Learner's Dictionary definition of 'cup'

cup /kʌp/
noun, verb
- *noun*
  1. [C] a small container shaped like a bowl, usually with a handle, used for drinking tea, coffee, etc.:
     - *a teacup*
     - *a coffee cup*
     - *a cup and saucer*
     - *a paper cup*
  2. [C] the contents of a cup:
     - *She drank the whole cup.*
     - *Would you like a cup of tea?*
  3. [C] a unit for measuring quantity used in cooking in the US; a metal or plastic container used to measure this quantity:
     - *two cups of flour and half a cup of butter*
  4. [C] a thing shaped like a cup:
     - *an egg cup*
  5. [C] a gold or silver cup on a STEM, often with two handles, that is given as a prize in a competition:
     - *She's won several cups for skating.*
     - *He lifted the cup for the fifth time this year.*
  6. [sing.] (usually Cup) a sports competition in which a cup is given as a prize:
     - *the World Cup*
  7. [C] one of the two parts of a BRA that cover the breast:
     - *a C cup*
  8. [C, U] a drink made from wine mixed with, for example, fruit juice
  9. [C] (NAmE) (in GOLF) a hollow in the ground that you must get the ball into
  10. [C] (NAmE) a piece of plastic that a man wears over his sex organs to protect them while he is playing a sport
      -more at SLIP *n*.
- *verb* (**-pp-**) [VN]
  1. **~ your hand(s) (around/over sth)** to make your hands into the shape of a bowl:
     - *She held the bird gently in cupped hands.*
  2. **~ sth (in your hands)** to hold sth, making your hands into a round shape:
     - *He cupped her face in his hands and kissed her.*

Appendix 3b: Oxford Advanced Learner's Dictionary definition of 'mug'

mug /mʌg/
  noun, verb
  ▪ *noun*
      1. A tall cup for drinking from, usually with straight sides and handle, used without a SAUCER:
          ◆ *a coffee mug*
          ◆ *a beer mug (= a large glass with a handle)*
      2. A mug and what it contains:
          ◆ *a mug of coffee*
      3. (slang) a person's face:
          ◆ *I never want to see his ugly mug again.*
      4. (informal) a person who is stupid and easy to trick:
          ◆ *They made me look a complete mug.*
          ◆ *He's no mug.*
  ▪ *verb* **(-gg-)**
      1. [VN] to attack sb violently in order to steal their money, especially in a public place:
          ◆ *She had been mugged in the street in broad daylight.*
      2. [V] ~ (for sb/sth) (informal, especially NAmE) to make silly expressions with your face or behave in a silly exaggerated way, especially on the stage or before a camera:
          ◆ *to mug for the cameras*

Appendix 4: Types of corpora

Type	Description	Function	Examples[1]
General/Reference	A corpus comprising texts from a variety of types or genres	Used as a reference or baseline for investigations of vocabulary, grammar or discourse structure, or to produce reference materials	BNC COBUILD
Monitor	A corpus comprising texts of the same type continually added to in the same proportions	Used to track changing patterns in a language	BoE
Specialised	A corpus comprising texts of a particular type	Designed to be representative of a given type of text with particular research projects in mind, eg. child language acquisition	CANCODE CHILDES
Comparable	Comprising two (or more) corpora in different languages or varieties of the same corpora	Used particularly by translators or learners to investigate differences and equivalencies between languages	ICE
Parallel	Comprising two (or more) of the same texts translated, eg. EU regulations	Used by translators or governments in areas with more than one official language	
Learner	Comprising a collection of texts produced by learners of a language	Used to identify how learners differ from each other and from native speakers	ICLE LOCNES LCLE
Historic/Diachronic	Comprising texts from different periods of time	Used to trace the development of language over time	Helsinki
Spoken	Comprising spoken text from various categories	Used most particularly for detailed prosodic studies	LLC SEC CSAE

[1] Key: BNC – British National Corpus
   COBUILD – Collins Birmingham University International Language Database
   BoE – Bank of English
   CANCODE – Cambridge and Nottingham Corpus of Discourse in English
   CHILDES – Child Language Data Exchange System
   ICE – International Corpus of English
   ICLE – Longman Corpus of Learners' English
   LOCNES – Louvain Corpus of Native English Essays
   Helsinki – Helsinki Corpus of English Texts: Diachronic Part
   LLC – London-Lund Corpus
   SEC – Lancaster/IBM Spoken English Corpus
   CSAE – Corpus of Spoken American English

(Hunston, 2002; Kennedy, 1998)